JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER

p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 2 September 2026
Hal: 1645-1653
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i2.1422

Optimasi Metode Naïve Bayes dan Random Forest Terhadap Akurasi Prediksi Gadget Addiction pada Anak-Anak

Nathania Asyifa[1*], Muhammad Syahputra Novelan[2]

[1,2]Pascasarjana, Magister Teknologi Informasi, Universitas Pembangunan Panca Budi, Medan, Indonesia
Email: [1*]nathaniasyifa09@gmail.com, [2]putranovelan@dosen.pancabudi.ac.id
(* Email Corresponding Author:nathaniasyifa09@gmail.com)
Received: 1 Juli 2026 | Revision: 8 Juli 2026 | Accepted: 11 Juli 2026

Abstrak

Penggunaan gagdet yang berlebihan pada anak-anak telah menjadi permasalahan kesehatan mental yang semakin mengkhawatirkan, dengan dampak negatif pada aspek psikologis, fisik, dan sosial. Deteksi dini terhadap perilaku kecanduan gadget (*gadget addiction*) sangat diperlukan agar intervensi dapat dilakukan secara tepat dan efesien. Penelitian ini bertujuan untuk membandingkan dan mengoptimasi kinerja dua algortima *machine learning*, yaitu Naïve Bayes dan *Random Forest*, dalam memprediksi tingkat kecanduan gadget pada anak-anak usia 5-12 tahun. Data dikumpulkan melalui kuesioner berbasis skala likert yang disebarkan kepada 150 responden (orangtua siswa) di SD Muhammadiyah 18 Medan, mencakup variabel durasi penggunaan gadget, gejala fisik, tingkat ketergantungan, serta faktor psikologis. Data diproses melalui tahapan *preprocessing*, kemudian dibagi 80:20 untuk data latih dan uji. Hasil evaluasi menujukkan bahwa algoritma Naive Bayes menghasilkan akurasi sebesar 95%, sedangkan *Random Forest* dengan paramater default maupun setelah optimasi mencapai akurasi 100% dengan nilai *presicion*, *recall*, dan *f1-Score* sempurna pada seluruh kelas. Perbedaan akurasi sebesar 5% antara kedua algoritma membuktikan bahwa Random Forest secara signifikan lebih unggul dibandingkan Naïve Bayes. Berdasarkan analisis feature importance, durasi penggunaan gadget menjadi fitur paling berpengaruh, sehingga solusi intervensi berupa pembatasan waktu layar dan strategi pengelolaan perilaku direkomendasikan sesuai tingkat kecanduan anak.

Kata Kunci: Kecanduan Gadget, *machine learning*, Naïve Bayes, *Random Forest*, anak-anak

Abstract

Excessive gadget use in children has become an increasingly worrying mental health problem, with negative impacts on psychological, physical, and social aspects. Early detection of gadget addiction behavior is essential for appropriate and efficient intervention. This study aims to compare and optimize the performance of two machine learning algorithms, namely Naïve Bayes and Random Forest, in predicting the level of gadget addiction in children aged 5–12 years. Data were collected through a Likert-based questionnaire distributed to 150 respondents (parents of students) at Muhammadiyah 18 Elementary School, Medan, covering variables such as gadget usage duration, physical symptoms, level of dependence, and psychological factors. The data were processed through a preprocessing stage, then divided with a ratio of 80:20 for training and testing data. The evaluation results showed that the Naïve Bayes algorithm produced an accuracy of 95%, while Random Forest with default parameters and after hyperparameter optimization achieved 100% accuracy with perfect precision, recall, and F1-score values across all classes. The 5% difference in accuracy between the two algorithms demonstrates that Random Forest is significantly superior to Naïve Bayes. Based on feature importance analysis, screen time is the most influential factor, so interventions such as screen time restrictions and behavioral management strategies are recommended based on the child's level of addiction..

Keywords*: Gadget Addiction, machine learning, Naïve Bayes, Random Forest, children*

1. PENDAHULUAN

Kemajuan teknologi digital dalam beberapa dekade terakhir telah membawa perubahan besar terhadap pola kehidupan anak-anak, terutama dalam hal interaksi dengan perangkat digital atau gadget. Smartphone, tablet, dan perangkat sejenisnya kini telah menjadi bagian tak terpisahkan dari keseharian anak-anak, bahkan pada kelompok usia yang sangat muda. Meskipun gadget memberikan manfaat dalam aspek pendidikan, hiburan, dan komunikasi, penggunaan yang tidak terkontrol berpotensi menimbulkan dampak serius berupa ketergantungan atau yang dikenal sebagai kecanduan gadget *(gadget addiction)*[1]. Kecanduan gadget pada anak-anak didefiniskan sebagai perilaku penggunaan perangkat digital secara kompulsif, yang ditandai dengan kesulitan untuk menghentikan penggunaan, rasa tidak nyaman ketikan menggunakan perangkat, serta ganggian terhadap aktivitas dan fungsi sehari-hari[2].

Dampak kecanduan gadget pada anak-anak bersifat multidimensional, mencakup aspek psikologis, fisik, maupun sosial. Secara psikologis, kondisi ini dapat memicu insomnia, perubahan perilaku sosial, rendahnya kepercayaan diri, serta kecemasan[3]. Penelitian di Indonesia mengungkapkan bahwa anak-anak yang menghabiskan waktu menggunakan gadget selama enam jam atau lebih per hari mengalami masalah emosional dan pola makan yang tidak sehat [4]. Dari sisi sosial, fenomena phubbing, yaitu perilaku mengabaikan orang lain di lingkungan sosial karena terlalu

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 2 September 2026
Hal: 1645-1653
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i2.1422

fokus pada perangkat, menjadi indikasi nyata bagaimana gadget dapat merusak kualitas interaksi sosial anak secara langsung [5]. Selain itu, penelitian juga mengidentifikasi adanya hubungan antara kecanduan gadget dengan gangguan emosional dan kemunculan tanda-tanda skizofrenia pada anak-anak [6]. Fakta-fakta ini menegaskan bahwa kecanduan gadget merupakan isu kesehatan mental yang mendesak dan memerlukan perhatian serius.

Mengingat kompleksitas dan luasnya dampak dari kecanduan gadget, deteksi dini dan prediksi yang akurat menjadi sangat krusial. Kemampuan untuk mengidentifikasi individu yang rentan terhadap kecanduan gadget sejak dini memungkinkan dilakukannya intervensi yang lebih cepat dan tepat sasaran, baik oleh orang tua, pendidik, maupun pemangku kebijakan. Namun demikian, proses deteksi manual yang selama ini banyak dilakukan cenderung bersifat subjektif dan membutuhkan waktu yang tidak sedikit. Oleh karena itu, diperlukan pendekatan yang lebih objektif, efisien, dan berbasis data, yakni dengan memanfaatkan teknologi kecerdasan buatan (Artificial Intelligence/AI) dan pembelajaran mesin (machine learning). Berbagai penelitian telah membuktikan bahwa machine learning mampu memprediksi berbagai kondisi kesehatan mental dengan tingkat akurasi yang tinggi [7], sehingga pendekatan ini layak diterapkan dalam konteks prediksi kecanduan gadget pada anak-anak.

Penelitian ini menawarkan solusi dengan menerapkan dan membandingkan dua algoritma machine learning yang telah terbukti efektif dalam berbagai kasus klasifikasi data, yaitu Naive Bayes dan Random Forest. Algoritma Naive Bayes dipilih karena kesederhanaannya, kecepatan komputasi, dan efisiensinya dalam menangani data berlabel serta kemampuannya bekerja dengan baik pada berbagai jenis data, termasuk data perilaku pengguna [8]. Sementara itu, algoritma Random Forest dipilih karena kemampuannya menangani data kompleks, ketangguhannya terhadap overfitting, serta kemampuannya mengidentifikasi fitur-fitur paling berpengaruh melalui analisis feature importance [9]. Dengan membandingkan dan mengoptimasi kedua metode ini melalui teknik hyperparameter tuning, diharapkan dapat ditemukan model prediksi yang paling efektif dan efisien untuk mendeteksi kecanduan gadget pada anak-anak berusia 5-12 tahun.
Dewi et al. [10] melakukan penelitian mengenai klasifikasi kecanduan smartphone pada pelajar sekolah menengah atas menggunakan metode machine learning berbasis feature weighting. Penelitian ini membandingkan algoritma C4.5, Naive Bayes (NB), dan K-Nearest Neighbor (KNN) dengan menerapkan seleksi atribut menggunakan Forward Selection, Backward Elimination, dan pendekatan korelasi Linear Regression. Hasil penelitian menunjukkan bahwa KNN dengan seleksi atribut menggunakan Linear Regression merupakan metode terbaik dengan akurasi 79,03%. Variabel yang paling berpengaruh meliputi usia, durasi penggunaan smartphone, aktivitas, dan kualitas tidur. Penelitian ini relevan karena secara langsung menguji Naive Bayes dalam konteks kecanduan smartphone, namun tidak melibatkan Random Forest sebagai metode pembanding maupun melakukan optimasi hyperparameter.Fitri et al. [11] meneliti penggunaan Random Forest dalam sistem klasifikasi kecemasan pada Generasi Z di Indonesia menggunakan data tweet yang diekstrak dari media sosial melalui teknik scraping. Model yang dikembangkan berhasil mencapai akurasi sangat tinggi, yakni 97,71%, dalam mengidentifikasi status kecemasan. Penelitian ini membuktikan keunggulan Random Forest dalam mengklasifikasikan masalah kesehatan mental berbasis data digital. Namun, penelitian ini hanya berfokus pada kecemasan pada remaja Generasi Z dan tidak membandingkan kinerjanya dengan algoritma probabilistik seperti Naive Bayes dalam konteks kecanduan perangkat digital pada anak-anak usia 5-12 tahun.

Parinduri et al. [12] menggunakan algoritma Naive Bayes untuk mengklasifikasikan tingkat kecanduan internet pada mahasiswa di Indonesia. Data dikumpulkan melalui kuesioner dan berhasil mengklasifikasikan level adiksi ke dalam beberapa kategori sehingga dapat menjadi dasar intervensi dan kebijakan penggunaan internet di institusi pendidikan. Penelitian ini sejenis dengan penelitian yang sedang dilakukan, namun memiliki perbedaan pada subjek (mahasiswa, bukan anak-anak), jenis kecanduan (internet, bukan gadget secara umum), serta tidak melakukan perbandingan dengan algoritma lain seperti Random Forest untuk mengukur performa yang lebih komprehensif.

Pania et al. [13] mengembangkan sistem berbasis website untuk mengklasifikasikan kecanduan bermain game online pada remaja dengan menggunakan metode Naive Bayes Classifier. Dengan 150 sampel data dan berbagai rasio training/testing, penelitian ini mencapai akurasi tertinggi sebesar 93%. Meskipun menggunakan algoritma yang sama dengan salah satu metode dalam penelitian ini, penelitian tersebut tidak melakukan optimasi hyperparameter maupun perbandingan dengan algoritma ensemble learning, serta berfokus pada kecanduan game online bukan kecanduan gadget secara umum pada anak-anak usia dini.

Fikri et al. [14] meneliti klasifikasi tingkat kecanduan internet pada remaja di Pekanbaru menggunakan algoritma Naive Bayes dengan data kuesioner sebanyak 510 responden. Hasil 10-fold cross-validation menunjukkan rata-rata akurasi sebesar 93%, dengan precision ~87,3%, recall ~89,9%, dan F1-score ~88,1%. Penelitian ini mempertegas efektivitas Naive Bayes dalam mengklasifikasikan kecanduan digital, namun juga memiliki keterbatasan yang sama, yakni hanya menggunakan satu algoritma tanpa membandingkannya dengan pendekatan ensemble seperti Random Forest, serta tidak melakukan optimasi model untuk meningkatkan performa.

Anshori et al. [3] menerapkan Backpropagation Neural Network (BPNN) untuk memprediksi kecanduan smartphone pada remaja dan berhasil mencapai akurasi sebesar 99,49%. Penelitian ini menunjukkan potensi besar algoritma berbasis kecerdasan buatan dalam mendeteksi kecanduan perangkat digital. Meski demikian, pendekatan neural network memerlukan sumber daya komputasi yang jauh lebih besar dan interpretasi hasilnya lebih sulit dibandingkan algoritma klasifikasi seperti Naive Bayes dan Random Forest. Penelitian ini juga tidak membandingkan kedua algoritma tersebut secara langsung dalam konteks prediksi kecanduan gadget pada anak-anak usia dini.

Ernawati et.al. [15] melakukan penelitian penerapan data mining untuk klasifikasi penduduk miskin menggunakan Random Forest dan K-Nearest Neighbors (KNN). Hasil penelitian tersebut menunjukkan bahwa Random

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 2 September 2026
Hal: 1645-1653
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i2.1422

Forest memberikan performa yang lebih baik dibandingkan KNN dengan nilai akurasi sebesar 0.6023, presisi 0.4827, recall 0.4177, F1-score 0.4479, dan AUC 0.5681. Temuan ini memperkuat posisi Random Forest sebagai algoritma yang andal dalam menangani klasifikasi data multidimensi.

Hamzah et.al. [16] dalam penelitiannya menggunakan Decision Tree C4.5 dan Naive Bayes untuk menganalisis indikasi penyakit diabetes menggunakan Rapid Miner. Penelitian tersebut menunjukkan bahwa algoritma Naive Bayes mampu melakukan klasifikasi dengan tingkat akurasi yang kompetitif dibandingkan metode lain, sehingga pendekatan ini layak untuk dikembangkan lebih lanjut pada permasalahan prediksi kesehatan dan perilaku manusia

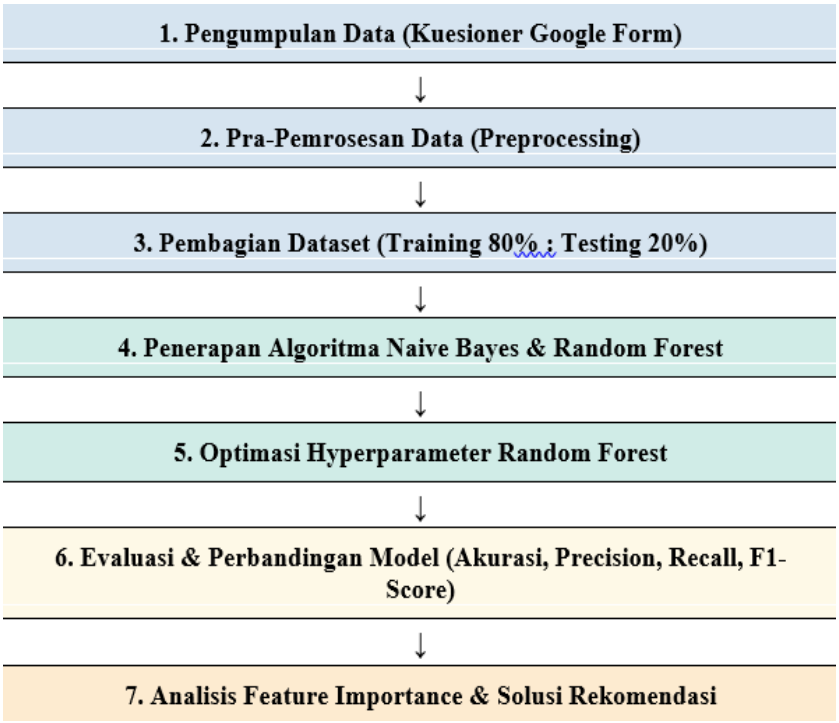
Berdasarkan tinjauan terhadap enam penelitian terkait di atas, terdapat kesenjangan (research gap) yang signifikan dalam literatur yang ada. Pertama, sebagian besar penelitian hanya menggunakan satu algoritma machine learning tanpa melakukan perbandingan komprehensif antara algoritma probabilistik dan ensemble learning. Kedua, penelitian yang secara khusus membandingkan dan mengoptimasi algoritma Naive Bayes dan Random Forest dalam konteks prediksi kecanduan gadget pada anak-anak usia 5-12 tahun masih sangat terbatas. Ketiga, sebagian besar penelitian berfokus pada remaja atau mahasiswa, sehingga penelitian pada kelompok anak-anak yang berada pada tahap perkembangan psikososial yang lebih rentan masih kurang mendapat perhatian. Keempat, tidak banyak penelitian yang menyertakan proses optimasi hyperparameter secara sistematis untuk meningkatkan performa model. Penelitian ini hadir untuk mengisi kesenjangan tersebut dengan melakukan perbandingan dan optimasi kedua algoritma secara langsung pada dataset kecanduan gadget anak-anak.

Berdasarkan latar belakang dan gap analysis yang telah diuraikan, penelitian ini bertujuan untuk: (1) mengukur dan membandingkan kinerja prediktif algoritma Naive Bayes dan Random Forest dalam mengklasifikasikan tingkat kecanduan gadget pada anak-anak usia 5-12 tahun; (2) melakukan optimasi hyperparameter pada algoritma Random Forest guna mendapatkan konfigurasi terbaik; (3) mengidentifikasi fitur-fitur yang paling berpengaruh terhadap prediksi kecanduan gadget melalui analisis feature importance; serta (4) merumuskan solusi intervensi berbasis data bagi orang tua dan pendidik berdasarkan hasil prediksi model. Melalui penelitian ini, diharapkan dapat tersedia suatu model machine learning yang akurat dan efisien sebagai alat deteksi dini kecanduan gadget pada anak-anak, sehingga upaya pencegahan dan penanganan dapat dilakukan secara lebih proaktif, tepat sasaran, dan berbasis bukti ilmiah (evidence-based) di era digital ini.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini merupakan penelitian kuantitatif yang menggunakan pendekatan komparatif berbasis machine learning untuk memprediksi tingkat kecanduan gadget pada anak-anak usia 5-12 tahun. Secara keseluruhan, penelitian dilaksanakan melalui enam tahapan sistematis yang saling berkaitan, mulai dari pengumpulan data hingga penarikan kesimpulan dan rekomendasi solusi. Alur kera penelitian ini digambarkan pada Gambar 1 berikut.



Gambar 1. Alur Tahapan Penelitian

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 2 September 2026
Hal: 1645-1653
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i2.1422

Tahap pertama adalah pengumpulan data yang dilakukan dengan menyebarkan kuesioner berbasis skala Likert (1–6) melalui platform Google Form kepada orang tua siswa SD Muhammadiyah 18 Medan. Total responden yang berhasil dikumpulkan sebanyak 150 orang tua dengan anak berusia 5–12 tahun. Kuesioner terdiri dari dua bagian, yaitu data demografis (jenis kelamin, usia, dan rata-rata durasi penggunaan gadget per hari) dan 10 pernyataan skala Likert mengenai perilaku dan dampak penggunaan gadget [2].

Tahap kedua adalah pra-pemrosesan data (preprocessing), yang mencakup tiga sub-tahap: (a) pembersihan data (data cleaning) berupa penghapusan data kosong dan duplikasi; (b) transformasi data berupa encoding variabel kategorikal dan normalisasi nilai numerik; serta (c) pelabelan (labeling) hasil kuesioner ke dalam tiga kelas target, yaitu Tidak Kecanduan, Cukup Kecanduan, dan Kecanduan Parah. Proses pelabelan dilakukan berdasarkan skor total kuesioner dengan rentang nilai yang telah ditentukan berdasarkan studi literatur terkait [6].

Tahap ketiga adalah pembagian dataset dengan rasio 80:20, di mana 80% data digunakan untuk pelatihan model (training set) dan 20% data digunakan untuk pengujian model (testing set). Pembagian ini dilakukan secara acak dengan teknik stratified splitting untuk memastikan distribusi kelas yang proporsional pada kedua subset data [10].

Tahap keempat adalah penerapan algoritma, yaitu implementasi Naive Bayes dan Random Forest menggunakan bahasa pemrograman Python dengan pustaka scikit-learn. Kedua model dilatih pada training set yang sama untuk memastikan perbandingan yang adil dan objektif. Tahap kelima adalah optimasi hyperparameter yang diterapkan khusus pada Random Forest menggunakan metode Grid Search Cross-Validation (GridSearchCV) dengan k-fold = 5. Tahap keenam adalah evaluasi dan perbandingan model menggunakan empat metrik, yaitu akurasi, precision, recall, dan F1-Score. Terakhir, tahap ketujuh adalah analisis feature importance dan perumusan solusi berbasis data.

2.2 Metode Penyelesaian Masalah

Bagian ini menjelaskan secara rinci metode penyelesaian masalah yang diterapkan dalam penelitian, meliputi algoritma Naive Bayes, algoritma Random Forest, tahapan optimasi hyperparameter, dan metrik evaluasi yang digunakan.

a. Spesifikasi Variabel Penelitian

Variabel penelitian terdiri dari variabel independen (X) dan variabel dependen (Y). Variabel independen mencakup seluruh atribut yang diperoleh dari hasil kuesioner, sedangkan variabel dependen merupakan label kelas kecanduan gadget. Rincian variabel disajikan pada Tabel 1 berikut.

Tabel 1. Variabel Penelitian

No.	Variabel	Jenis	Keterangan
1	Jenis Kelamin	Independen (X)	Kategorikal
2	Usia Anak	Independen (X)	Numerik (5-12 tahun)
3	Durasi Penggunaan Gadget/Hari	Independen (X)	Numerik (jam)
4	Tujuan Penggunaan Gadget	Independen (X)	Kategorikal
5	Q1 s.d. Q10 (Skala Likert)	Independen (X)	Ordinal (1-6)
6	Tingkat Kecanduan Gadget	Dependen (Y)	3 Kelas (Tidak, Cukup, Parah)

b. Algoritma Naïve Bayes

Naive Bayes (NB) adalah algoritma klasifikasi probabilistik yang bekerja berdasarkan Teorema Bayes dengan asumsi bahwa setiap fitur bersifat independen satu sama lain [8]. Algoritma ini dipilih sebagai baseline model dalam penelitian ini karena kesederhanaan implementasinya dan kemampuannya bekerja secara efisien pada dataset berdimensi tinggi. Probabilitas posterior dihitung menggunakan rumus pada persamaan (1) berikut.

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)} \quad (1)$$

Keterangan: P(C|X) adalah probabilitas data X termasuk ke dalam kelas C (posterior); P(X|C) adalah probabilitas kemunculan fitur X pada kelas C (likelihood); P(C) adalah probabilitas awal kelas C (prior); dan P(X) adalah probabilitas total fitur X (evidence). Dalam klasifikasi multivariat, likelihood dihitung sebagai perkalian semua probabilitas fitur seperti pada persamaan (2).

$$P(X|C) = \prod_{i=1}^{n} P(xi_i|C) \quad (2)$$

Keterangan: X = (x1, x2, ..., xn) adalah kumpulan fitur; P(xi|C) adalah probabilitas fitur ke-i pada kelas tertentu; dan simbol $\prod$ menyatakan perkalian seluruh probabilitas fitur. Asumsi independensi ini dikenal sebagai naive assumption, dan meskipun jarang terpenuhi sepenuhnya dalam data nyata, Naive Bayes tetap menunjukkan kinerja yang kompetitif dalam berbagai kasus klasifikasi [12].

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 2 September 2026
Hal: 1645-1653
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i2.1422

c. Algoritma Random Forest

Random Forest (RF) adalah algoritma ensemble learning yang membentuk sekumpulan pohon keputusan (decision trees) dari berbagai subset data dan fitur secara acak, kemudian menggabungkan prediksi masing-masing pohon melalui proses majority voting [9]. Prediksi akhir dihasilkan berdasarkan persamaan (3) berikut.

$$\hat{y} = mode(h1_1(x), h2_2(x), \ldots, hk_k(x)) \quad (3)$$

Keterangan: ŷ adalah prediksi akhir; h1(x), h2(x), ..., hk(x) adalah prediksi dari masing-masing pohon keputusan ke-1 hingga ke-k; dan mode menyatakan nilai yang paling sering dipilih oleh seluruh pohon. Karena setiap pohon dilatih dengan data bootstrap sampling yang berbeda dan pemilihan fitur secara acak, Random Forest memiliki ketahanan tinggi terhadap overfitting dan mampu menangkap pola non-linier yang kompleks [11].

d. Optimasi Random Forest

Untuk memaksimalkan kinerja Random Forest, dilakukan optimasi hyperparameter menggunakan metode Grid Search Cross-Validation (GridSearchCV) dengan k-fold = 5. Ruang pencarian hyperparameter yang diuji disajikan pada Tabel 2 berikut.

Tabel 2. Ruang Pencarian Hyperparameter Random Forest

Hyperparameter	Nilai yang Diuji	Deskripsi
n_estimators	50, 100, 200	Jumlah pohon keputusan dalam hutan
max_depth	None, 5, 10	Kedalaman maksimum setiap pohon
min_samples_split	2, 5, 10	Minimum sampel untuk membagi node

Proses Grid Search menguji semua kombinasi hyperparameter yang mungkin (total 3 × 3 × 3 = 27 kombinasi) dan memilih kombinasi yang menghasilkan akurasi validasi tertinggi. Pendekatan ini memastikan bahwa konfigurasi hyperparameter yang dipilih bukan sekadar asumsi, melainkan hasil pengujian empiris yang tervalidasi secara silang [10].

e. Metrik Evaluasi Model

Kinerja kedua algoritma dievaluasi menggunakan empat metrik utama. Metrik-metrik ini dipilih karena mampu memberikan gambaran kinerja model yang komprehensif, khususnya pada dataset yang berpotensi tidak seimbang (imbalanced dataset) [8]. Keempat metrik tersebut adalah sebagai berikut.

Akurasi mengukur proporsi prediksi yang benar terhadap total keseluruhan prediksi, sebagaimana ditunjukkan pada persamaan (4).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

Precision mengukur seberapa akurat model ketika memprediksi kelas positif, dihitung berdasarkan persamaan (5).

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

Recall mengukur kemampuan model dalam mendeteksi seluruh data positif yang sebenarnya, dihitung berdasarkan persamaan (6).

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

F1-Score merupakan harmonic mean antara precision dan recall, digunakan ketika keseimbangan antara keduanya penting. Rumus F1-Score ditunjukkan pada persamaan (7).

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (7)$$

Keterangan untuk persamaan (4) hingga (7): TP (True Positive) adalah prediksi benar untuk kelas positif; TN (True Negative) adalah prediksi benar untuk kelas negatif; FP (False Positive) adalah prediksi salah yang mengklasifikasikan data negatif sebagai positif; dan FN (False Negative) adalah prediksi salah yang mengklasifikasikan data positif sebagai negatif [8].

3. HASIL DAN PEMBAHASAN

3.1 Deskripsi Data Penelitian

Tahapan awal sebelum penerapan model machine learning adalah memahami karakteristik data yang dikumpulkan. Data penelitian ini bersumber dari kuesioner yang disebarkan kepada 150 orang tua siswa SD Muhammadiyah 18 Medan

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 2 September 2026
Hal: 1645-1653
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i2.1422

yang memiliki anak berusia 5–12 tahun. Proses pengumpulan data berlangsung dari November 2025 hingga Februari 2026. Bagian berikut menyajikan distribusi data berdasarkan aspek demografis dan hasil preprocessing.

3.2 Implementasi dan Hasil Penerapan Model (bila ada)

Bagian ini menyajikan hasil implementasi dua algoritma machine learning yang diterapkan pada dataset penelitian, meliputi kinerja model Naïve Bayes sebagai baseline, kinerja Random Forest dengan paramater default, dan hasil setelah optimasi hyperparameter. Seluruh eksperimen dilakukan menggunakan bahasa pemrograman python dengan dataset 80:20.

3.2.1 Hasil Model Naïve Bayes

Algoritma Naïve Bayes diterapkan sebagai model pembanding (baseline). Model dilatih pada data latih kemudian dievaluasi pada data uji.Hasil evaluasi model Naïve Bayes pada data uji ditampilkan secara lengkap pada Tabel 3 berikut, dengan perincian per kelas beserta metrik evaluasinya.

Tabel 3. Hasil Evaluasi Model Naïve Bayes

Kelas	Precision	Recall	F1-Score
Tidak Kecanduan	94%	94%	94%
Cukup Kecanduan	92%	100%	96%
Kecanduan Parah	100%	92%	96%
Akurasi	—	—	**95%**
Macro Avg	95%	95%	95%
Weighted Avg	95%	95%	95%

Bedasarkan tabel 5, model Naïve Bayes berhasil mencapai akurasi keseluruhan sebesar 95% pada data uji.

3.2.2 Hasil Model Random Forest

Algoritma Random Forest diimplementasikan dengan konfigurasi parameter default dari pustaka scikit-learn, yaitu n_estimators=100, max_depth=None, dan min_samples_split=2. Hasil evaluasi model Random Forest dengan parameter default disajikan pada Tabel 4.

Tabel 4. Hasil Evaluasi Model Random Forest

Kelas	Precision	Recall	F1-Score
Tidak Kecanduan	100%	100%	100%
Cukup Kecanduan	100%	100%	100%
Kecanduan Parah	100%	100%	100%
Akurasi	—	—	**100%**
Macro Avg	100%	100%	100%
Weighted Avg	100%	100%	100%

Tabel 6 memperlihatkan bahwa model Random Forest dengan parameter default berhasil mencapai akurasi sempurna sebesar 100% pada seluruh kelas data uji.

3.2.3 Hasil Optimasi Random Forest

Meskipun Random Forest dengan parameter default telah mencapai akurasi 100%, proses optimasi hyperparameter tetap dilakukan untuk memvalidasi apakah konfigurasi parameter default memang merupakan yang optimal untuk dataset ini.

Hasil ini mengindikasikan bahwa karakteristik dataset penelitian memiliki pola yang cukup jelas dan terstruktur sehingga dapat dipelajari secara sempurna oleh Random Forest bahkan dengan konfigurasi standar. Tabel 5 menyajikan perbandingan performa antara ketiga model yang diuji.

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 2 September 2026
Hal: 1645-1653
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i2.1422

Tabel 5. Perbandingan Performa Ketiga Model

Model	Akurasi	Precision	Recall	F1-Score
Naive Bayes	95%	95%	95%	95%
Random Forest (Default)	**100%**	**100%**	**100%**	**100%**
Random Forest (Teroptimasi)	**100%**	**100%**	**100%**	**100%**

3.3 Uji Hipotesis dan Pembahasan

Penelitian ini mengajukan empat hipotesis tambahan yang diuji berdasarkan hasil perbandingan metrik evaluasi. Rangkuman hasil uji hipotesis disajikan pada Tabel 6 berikut.

Tabel 6. Hasil Uji Hipotesis Penelitian

Hipotesis	Pernyataan	Selisih	Status
H_1a	RF memiliki akurasi lebih tinggi dari NB	+5%	**DITERIMA**
H_1b	RF memiliki precision lebih tinggi dari NB	+5%	**DITERIMA**
H_1c	RF memiliki recall lebih baik dari NB	+5%	**DITERIMA**
H_1d	RF menghasilkan F1-score lebih unggul dari NB	+5%	**DITERIMA**
H_0	Tidak ada perbedaan signifikan antara RF dan NB	—	**DITOLAK**

Berdasarkan Tabel 6, seluruh hipotesis alternatif (H_1a hingga H_1d) diterima, sementara hipotesis nol (H_0) ditolak. Random Forest terbukti secara konsisten lebih unggul dibandingkan Naive Bayes pada semua metrik evaluasi dengan perbedaan sebesar 5% untuk seluruh metrik. Temuan ini menunjukkan keunggulan Random Forest dalam klasifikasi masalah kesehatan mental, serta membuktikan superioritas Random Forest dibandingkan algoritma lain dalam kasus klasifikasi multidimensi.

Keunggulan Random Forest atas Naive Bayes dalam penelitian ini dapat dijelaskan melalui beberapa aspek berikut.

a. Kemampuan Menangkap Korelasi Antar Fitur: Naive Bayes mengasumsikan independensi antar fitur (naive assumption), padahal dalam data perilaku kecanduan gadget, fitur-fitur seperti durasi penggunaan dan tingkat ketergantungan cenderung saling berkorelasi. Random Forest tidak memiliki asumsi ini sehingga mampu menangkap hubungan tersebut lebih baik.
b. Mekanisme Ensemble: Dengan menggabungkan 100 pohon keputusan, Random Forest mereduksi varians prediksi secara signifikan, menghasilkan model yang lebih stabil dan generalisasi yang lebih kuat.
c. Ketahanan terhadap Noise: Random Forest lebih toleran terhadap nilai-nilai ekstrem dan data yang tidak representatif karena keputusan diambil secara kolektif melalui majority voting, bukan oleh satu model tunggal.

3.4 Solusi dan Rekomendasi Berbasis Data

Berdasarkan hasil analisis dan performa model, penelitian ini merumuskan solusi intervensi yang dapat diimplementasikan oleh orang tua dan pendidik berdasarkan kategori kecanduan yang terdeteksi. Solusi ini dirancang bersifat bertingkat (tiered intervention) sesuai tingkat keparahan kecanduan.

3.4.1 Rekomendasi Berdasarkan Tingkat Kecanduan

a. Anak Kategori Tidak Kecanduan (Skor rendah, durasi < 2 jam/hari): Pertahankan pola penggunaan dengan pendampingan orang tua. Kenalkan kegiatan non-digital yang menarik sebagai alternatif hiburan. Lakukan monitoring berkala setiap 3 bulan menggunakan kuesioner yang sama.
b. Anak Kategori Cukup Kecanduan (Skor sedang, durasi 2–4 jam/hari): Terapkan batasan durasi penggunaan 1,5 jam per hari dengan monitoring ketat. Jadwalkan waktu tanpa gadget minimal 2 jam per hari, terutama saat makan dan sebelum tidur. Gunakan aplikasi parental control untuk membatasi akses konten dan durasi.
c. Anak Kategori Kecanduan Parah (Skor tinggi, durasi > 4 jam/hari): Terapkan gadget detox bertahap dengan mengurangi durasi 30 menit per minggu hingga mencapai kurang dari 1 jam per hari. Pertimbangkan konsultasi dengan psikolog anak jika gejala kecemasan atau gangguan tidur muncul. Libatkan guru BK sekolah dalam program pemulihan.

3.4.2 Rekomendasi Berdasarkan Fitur Kritis

Berdasarkan empat fitur dengan importance score tertinggi, berikut adalah rekomendasi spesifik yang dapat diterapkan.

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 2 September 2026
Hal: 1645-1653
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i2.1422

a. Q1 – Prioritas Terganggu: Gunakan aplikasi pengunci otomatis yang membatasi akses gadget pada waktu-waktu kritis seperti jam belajar (16.00–18.00), waktu makan, dan 1 jam sebelum tidur (21.00–22.00).
b. Q3 – Dampak Fisik: Terapkan aturan 20-20-20: setiap 20 menit menggunakan layar, istirahat selama 20 detik dengan melihat objek sejauh 20 kaki (±6 meter). Pastikan postur duduk yang ergonomis dan pencahayaan yang cukup.
c. Q9 – Escapism: Ajarkan anak teknik regulasi emosi yang sehat tanpa gadget, seperti menceritakan perasaan kepada orang tua, aktivitas fisik ringan, atau hobi seni. Identifikasi sumber stres utama anak dan tangani secara langsung.
d. Q10 – Ketergantungan: Mulai terapkan "waktu bebas gadget" selama 30 menit per hari pada minggu pertama, kemudian tingkatkan secara bertahap hingga 2–3 jam per hari. Buat aktivitas pengganti yang menarik dan melibatkan interaksi sosial langsung.

Secara keseluruhan, penelitian ini menunjukkan bahwa implementasi machine learning khususnya Random Forest, sangat potensial sebagai tulang punggung sistem deteksi dini kecanduan gadget pada anak-anak. Model yang dihasilkan tidak hanya akurat dalam klasifikasi, tetapi juga informatif melalui analisis feature importance yang menghasilkan rekomendasi berbasis data. Keterbatasan utama penelitian ini adalah ukuran dataset yang masih terbatas (150 sampel) dan pengumpulan data yang hanya dilakukan di satu sekolah, sehingga generalisasi model ke populasi yang lebih luas masih perlu divalidasi dengan dataset yang lebih besar dan beragam.

4. KESIMPULAN

Penelitian ini telah berhasil menjawab seluruh rumusan masalah yang diajukan melalui perbandingan dan optimasi algoritma Naïve Bayes dan *Random Forest* dalam memprediksi tingkat kecanduan gadget pada anak-anak usia 5-12 tahun. Berdasarkan hasil evaluasi pada data uji, algoritma Naiva Bayes berhasil mencapai akurasi sebesar 95% dengan nilai *precission, recall*, dan *F1-Scrore* rata-rata sebesar 95%. Kinerja ini membuktikan bahwa Naïve Bayes dan layak dijadikan model baseline yang handal dalam konteks prediksi perilaku kecanduan gadget berbasis data kuesioner. Di sisi lain, algoritma *Random Forest* baik dengan konfigurasi paramatert default maupun setelah proses optimasi, berhasil mencapai akurasi sempurna sebesar 100% dengan nilai *precission*, *recall*, dan *F1-score* sebesar 100%. Selisish akurasi 5%antara kedua algoritma secara statistik membuktikan bahwa terdapat perbedaan kinerja yang signifikan. Hasil analisis mengungkapkan bahwa durasi penggunaan gadget per hari merupakan prediktor paling dominan (18,7%) dalam menentukan tingkat kecanduan, diikuti oleh indikator ketergantungan emosional seperti kebiasaan membawa gadget ke mana pun (Q10, 15,4%) dan penggunaan gadget sebagai mekanisme pelarian dari perasaan tidak nyaman (escapism) (Q9, 14,3%). Temuan ini memiliki implikasi praktis yang signifikan: intervensi yang efektif harus menyasar secara bersamaan aspek pembatasan waktu layar dan pengelolaan aspek psikologis anak, bukan hanya berfokus pada pembatasan akses semata. Berdasarkan hasil tersebut, penelitian ini merumuskan solusi intervensi bertingkat yang disesuaikan dengan kategori kecanduan, mulai dari monitoring berkala bagi anak yang tidak kecanduan, pembatasan durasi dengan parental control bagi kategori cukup kecanduan, hingga program gadget detox terstruktur dan rujukan ke psikolog anak untuk kategori kecanduan parah. Meskipun hasil penelitian ini sangat menjanjikan, terdapat beberapa keterbatasan yang perlu diakui. Dataset yang digunakan masih relatif terbatas dari satu sekolah, sehingga kemampuan generalisasi model ke populasi yang lebih luas belum dapat sepenuhnya dijamin. Selain itu, data dikumpulkan melalui kuesioner yang diisi oleh orang tua, sehingga rentan terhadap bias subjektif dalam pelaporan (self-reporting bias). Untuk penelitian selanjutnya, disarankan agar dilakukan pengumpulan data dengan sampel yang lebih besar dan beragam dari berbagai latar belakang demografis dan geografis, serta mengintegrasikan data objektif seperti log penggunaan aplikasi dari perangkat. Eksplorasi algoritma yang lebih canggih seperti XGBoost, LightGBM, atau arsitektur Deep Learning juga direkomendasikan sebagai perbandingan di masa mendatang. Secara keseluruhan, penelitian ini memberikan kontribusi nyata bagi pengembangan sistem berbasis kecerdasan buatan untuk mendukung kesejahteraan anak-anak di era digital, dan diharapkan dapat menjadi landasan ilmiah bagi orang tua, pendidik, dan pembuat kebijakan dalam merancang program pencegahan serta penanganan kecanduan gadget yang lebih efektif dan berbasis bukti.

REFERENCES

[1] R. G. Pratiwi dan R. U. Malwa, "*Faktor yang Mempengaruhi Kecanduan Gadget terhadap Perilaku Remaja*," *Jurnal Ilmiah Psyche*, vol. 15, no. 2, pp. 105–112, 2021. doi: 10.33557/jpsyche.v15i2.1550.

[2] E. M. Abu-Taieh, I. AlHadid, K. Kaabneh, R. S. Alkhawaldeh, S. Khwaldeh, R. Masa'deh, dan A. Alrowwad, "*Predictors of Smartphone Addiction and Social Isolation among Jordanian Children and Adolescents Using SEM and ML*," *Big Data and Cognitive Computing*, vol. 6, no. 3, pp. 92, 2022. doi: 10.3390/bdcc6030092.

[3] M. I. Anshori, M. S. Haris, dan W. T. Kusuma, "*Penerapan Backpropagation Neural Network (BPNN) untuk Prediksi Kecanduan Smartphone pada Remaja*," *CICES: Conference of Islamic Civilization and Engineering Science*, vol. 9, no. 2, pp. 103–109, 2023. [Online].

JURNAL KOMPUTER TEKNOLOGI INFORMASI SISTEM KOMPUTER
p-ISSN: 2962-3022 | e-ISSN: 2963-7104
Volume 5 No 2 September 2026
Hal: 1645-1653
http://ejurnal.lkpkaryaprima.id/index.php/juktisi
DOI: https://doi.org/10.62712/juktisi.v5i2.1422

[4] P. D. Untari, D. Q. Ayuni, dan M. F. Dielsa, "*Dampak Penggunaan Gadget terhadap Gangguan Emosi dan Status Gizi pada Remaja 11–14 Tahun*," *Jurnal Kesehatan*, vol. 12, no. Supplementary 1, pp. 367–371, 2021. [Online].

[5] N. Jannatuna'im dan Fikrie, "*Phubbing dan Interaksi Sosial Anak*," 2022. [Online].

[6] O. Okfalisa, E. Budianita, M. Irfan, H. Rusnedy, dan S. Saktioto, "*The Classification of Children Gadget Addiction: The Employment of Learning Vector Quantization 3*," *IT Journal Research and Development*, vol. 5, no. 2, pp. 158–170, 2021. doi: 10.25299/itjrd.2021.vol5(2).5681.

[7] S. Verma, S. Sharma, dan P. Arora, "*Mental Health Analysis with Artificial Intelligence*," *International Journal of Scientific Research in Science and Technology*, vol. 11, no. 2, pp. 1–6, 2025. [Online].

[8] N. Mardiah, L. Marlina, K. Khairul, Z. Sitorus, dan M. Iqbal, "*Analysis of Indonesian People's Sentiment Towards 2024 Presidential Candidates on Social Media Using Naïve Bayes Classifier and Support Vector Machine*," *Building of Informatics, Technology and Science (BITS)*, vol. 6, no. 2, pp. 950–960, 2024. [Online].

[9] N. D. Ariyanta, A. N. Handayani, J. T. Ardiansah, dan K. Arai, "*Ensemble Learning Approaches for Predicting Heart Failure Outcomes: A Comparative Analysis of Feedforward Neural Networks, Random Forest, and XGBoost*," *Applied Engineering and Technology*, vol. 3, no. 3, pp. 173–184, 2024. [Online].

[10] N. K. C. B. Dewi, N. K. A. Wirdiani, dan D. M. S. Arsa, "*Klasifikasi Kecanduan Smartphone pada Pelajar Sekolah Menengah Atas menggunakan Metode Machine Learning Berbasis Feature Weighting*," *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, vol. 8, no. 1, pp. 95–104, 2022. doi: 10.26418/jp.v8i1.51914.

[11] R. R. Fitri, Asriyanik, dan W. Apriandari, "*Penggunaan Random Forest dalam Sistem Klasifikasi Kecemasan pada Generasi Z*," *Jurnal Informatika dan Teknik Elektro Terapan (JITET)*, vol. 13, no. 3, pp. 561–568, 2025. doi: 10.23960/jitet.v13i3.6905.

[12] F. Z. Parinduri, R. Dewi, dan Susiani, "*Klasifikasi Tingkat Kecanduan Internet pada Mahasiswa menggunakan Algoritma Naïve Bayes*," *JOMLAI: Journal of Machine Learning and Artificial Intelligence*, vol. 1, no. 3, pp. 257–264, 2022. doi: 10.55123/jomlai.v1i3.965.

[13] T. S. Pania, R. Hidayati, dan Kasliono, "*Klasifikasi Kecanduan Bermain Game Online pada Remaja menggunakan Metode Naïve Bayes Classifier Berbasis Website*," *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 4, no. 5, pp. 2518–2526, 2024. doi: 10.30865/klik.v4i5.1782.

[14] M. I. Fikri, E. Budianita, I. Iskandar, dan E. P. Cynthia, "*Klasifikasi Tingkat Kecanduan Internet terhadap Remaja Pekanbaru melalui Pendekatan Algoritma Naïve Bayes*," *ZONAsi: Jurnal Sistem Informasi*, vol. 6, no. 2, pp. 424–431, 2024. doi: 10.31849/zn.v6i2.20191.

[15] I. Hamzah, Z. Sitorus, dan Khairul, "*Analisa Classification Decision Tree C45 dan Naïve Bayes Pada Indikasi Penyakit Diabetes Menggunakan Rapid Miner," Jurnal Nasional Teknologi Komputer*, vol. 4, no. 1, pp. 25–33, 2024. doi: 10.61306/jnastek.v4i1.126.

[16] A. Ernawati, Khairul, Z. Sitorus, M. Iqbal, dan D. Nasution, "*Penerapan Data Mining Untuk Klasifikasi Penduduk Miskin Di Kabupaten Labuhanbatu Menggunakan Random Forest Dan K-Nearest Neighbors," Bulletin of Information Technology (BIT)*, vol. 6, no. 2, pp. 23–35, 2025. doi: 10.47065/bit.v6i1.1783.